DOCUMENT RESUME

ED 472 897 IR 058 602

AUTHOR	Byrum, John D., Jr.
TITLE	Challenges of Providing Bibliographic Access to Remote Electronic Resources in National Bibliographies: Problems and Solutions--An Overview.
PUB DATE	2002-08-00
NOTE	8p.; In: Libraries for Life: Democracy, Diversity, Delivery. IFLA Council and General Conference: Conference Programme and Proceedings (68th, Glasgow, Scotland, August 18-24, 2002); see IR 058 549.
AVAILABLE FROM	For full text: http://www.ifla.org.
PUB TYPE	Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Access to Information; Bibliographic Databases; Bibliographic Records; *Bibliographies; Cataloging; Cooperative Programs; Information Industry; Information Services; Information Sources; International Cooperation; Metadata; National Programs; *Online Systems

ABSTRACT

The proliferation of online resources has resulted in creating and accentuating challenges for national bibliographic agencies throughout the world. Through strategic planning and innovative approaches, providers of national bibliographies are seeking to realize bibliographic control of remote electronic material. For success, they will need to achieve a variety of goals, including: pursuing cooperative database building, re-purposing bibliographic information, and adapting selection criteria and levels of cataloguing to the special features of the digital world to achieve effective coverage of these resources; developing more automated tools for creating and maintaining bibliographic information and metadata; fostering increased research and development to improve cataloguing tools; expanding educational and training opportunities to prepare cataloguers and other library staff to better understand and service electronic resources; and, establishing a wide array of collaborative ventures with partners from throughout the information industry to gain needed resources to make it possible to meet the challenges of electronic resources. (Contains 14 references.) (Author)

Reproductions supplied by EDRS are the best that can be made
from the original document.

IFLA GLASGOW 2002

68th IFLA Council and General Conference August 18-24, 2002

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. Koopman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Code Number: 117-152-E
Division Number: IV
Professional Group: Bibliography - Workshop
Joint Meeting with: National Libraries
Meeting Number: 152
Simultaneous Interpretation: -

ED 472 897

Challenges of providing bibliographic access to remote electronic resources in national bibliographies: Problems and solutions – an overview

John D. Byrum, Jr.
Regional and Cooperative Cataloging Division
Library of Congress
Washington, D. C.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ABSTRACT:

The proliferation of online resources has resulted in creating and accentuating challenges for national bibliographic agencies throughout the world. Through strategic planning and innovative approaches, providers of national bibliographies are seeking to realize bibliographic control of remote electronic material. For success, they will need to achieve a variety of goals, including: pursuing cooperative database building, re-purposing bibliographic information, and adapting selection criteria and levels of cataloguing to the special features of the digital world to achieve effective coverage of these resources; developing more automated tools for creating and maintaining bibliographic information and metadata; fostering increased research and development to improve cataloguing tools; expanding educational and training opportunities to prepare cataloguers and other library staff to better understand and service electronic resources; and, establishing a wide array of collaborative ventures with partners from throughout the information industry to gain needed resources to make it possible to meet the challenges of electronic resources.

IR058602

The emergence of electronic material as a type of informational resource of value to libraries and library users can be traced back to the 1960s, if not before. Indeed, electronic works had become of sufficient

BEST COPY AVAILABLE

interest to the profession by the 1970s that groups were formed to formulate cataloguing rules and practices by which machine-readable data files and programs could be described and accessed within the parameters of national bibliographies and library catalogues. In 1977,the *ISBD(NBM): International Standard Bibliographic Description for Non-Book Materials*[1] and this publication set out rules for description of electronic formats for the first time within the IFLA framework. Although the only electronic resource provided for by this standard was machine-readable data files, by 1990 computer files, defined as data files and programs, had been allocated their own ISBD.[2] But, it was during the decade which followed that the nature and availability of electronic resources proliferated, necessitating only seven years later a major overhaul of this ISBD. Thus, in 1997 the *International Standard Bibliographic Description for Electronic Resources*[3] was published, covering the entire array of available handheld and remote electronic resources. Clearly, IFLA has successfully provided leadership through its ISBD program in formulating standard bibliographic descriptions to cover the gamut of new formats emerging from the technological revolution that closed the 20th century and continues even today. The ISBD(ER) is widely applied by national and other libraries throughout the world, either directly or indirectly through incorporation into most if not all national and multi-national cataloguing codes. In addition, there are, of course, a variety of metadata schemes from which to select for simpler and briefer descriptions of Web material.

As a result, specification of descriptive requirements needed to establish bibliographic control of remote electronic resources is not a problem needing to be addressed. Nevertheless, the incredible proliferation of networked information in a variety of complex and often mutable formats presents challenges so numerous and difficult as to raise the possibility of bibliographic chaos. From the point of view of the national bibliographic agency, bearing responsibility of creating an ongoing record of cultural heritage, the first and perhaps greatest need is to establish strategic response to this proliferation. National libraries on their own initiative and in some cases encouraged by legal deposit requirements have come to accept responsibility to expand the scope of their national bibliographies to include appropriate networked resources. It may prove necessary in some cases for the national bibliographic agency to remind or convince its funding sources of the importance of Web materials as an essential component in the overall "national imprint."

Whether mandated or undertaken as an initiative by the national bibliography, the additional workload of listing Web resources will rarely be offset by sufficient additional staffing, so national bibliographic agencies should expect to be challenged to develop new approaches to enable improved bibliographic control of proliferating electronic materials. National bibliographic agencies are finding that traditional policies and practices based on print and other tangible formats do not serve them adequately when confronting remote electronic resources.

Establishing the scope or coverage to which a national bibliographic agency will commit is the obvious early step in determining its strategic response. The "Final recommendations of the International Conference on National Bibliographic Services"[4] advises that "national bibliographies should include the current national output, and where practicable they should also provide retrospective coverage. When necessary, selection criteria should be defined and published by the national bibliographic agency." Because of the large and growing number of Web resources and the temporary nature of so many of them, it is quite likely that, as a practical matter, compilers of national bibliographies will need to adopt selection criteria to limit inclusion to materials that are of sufficient research or cultural value as to merit registration. In short, while national bibliographies might seek to include entries comprehensively for printed publications and perhaps also traditional non-book materials, they well may need to identify a sub-set of electronic resources to list.

Studies have revealed that Web resources are available for the full range of topics of interest to

researchers worldwide[5]. Traditional criteria often used by libraries, archives, and museums for printed publications and other non-book materials might assist national bibliographies in establishing criteria for inclusion of electronic resources. Such criteria include authorship, content, provenance, accuracy, relevance to institutional mission, and subject matter. In addition, remote electronic resources entail other characteristics to consider determining which of them justify listing in the national bibliography. Relevant characteristics include design of the resource, ease of use, timeliness of content, permanence, quality of links to other sites, value-added utility beyond print versions, scholarly reputation of the originating domain, uniqueness, and persistence of the resource itself and the URL by which it is accessed. It will likely be necessary to differentiate those resources that are freely available from those that are commercially, not only in terms of selecting and collecting them but also in terms of providing access to them.

National bibliographic services might well want to develop or encourage development of software to assist with resource selection and evaluation. What I have in mind are tools that would examine the characteristics of the resources, the extent and nature of their linkages, and use patterns, and report results according to specified criteria. This is but one example of how new automated products might facilitate the work of national libraries in coping with electronic resources. Later I shall mention a few more possibilities.

Another approach to providing coverage of electronic resources in our national bibliographies takes the direction of encouraging partnerships with other agencies that catalogue such materials. The bibliographic descriptions produced by these partners could serve as the basis of entries in national bibliographies, either directly contributed according to arrangements established by the cooperating institutions or indirectly derived from records residing in utilities or OPACs. In the U. S., for example, OCLC and RLG maintain bibliographic databases that are already rich in cataloguing for Web resources, and libraries everywhere, including the Library of Congress, re-use records from these databases for their catalogues and bibliographic products. In addition, there are many other agencies, such as government printing offices, to enlist in a coordinated effort to increase entries for electronic resources in the national bibliography. Essential to the success of such partnerships are clear agreements as to the cataloguing standards to be followed for the descriptions but also those pertinent to formulating standardized access points and, of course, use of MARC formats. The Program for Cooperative Cataloging from which the Library of Congress routinely garners data contributed by several hundred member libraries committed to uniform standards provides one model for national bibliographic services to consider when reaching out for help in increasing their coverage of remote electronic materials.

In many cases, those who manage national bibliographic services are already quite experienced in establishing partnerships with publishers, particularly those who produce print products. The challenges of dealing with digitized or born digital material will greatly expand the concept to include collaboration with metadata producers; standards developers; systems and software vendors; computing and technology suppliers; scholarly and academic enclaves; dot.com creators; bibliographic utilities; registration agencies; government agencies; other libraries and perhaps museums and archives as well.

The recommendations from the International Conference on National Bibliographic Services call on the national bibliographic agency to "take a lead in the updating and maintenance of national and international standards and principles" manifested in the records they produce and distribute.[6] As already pointed out, current standards for bibliographic descriptions for the full range of electronic resources fortunately exist in the form of the ISBD(ER). Nevertheless, there are left unaddressed several cataloguing problems that national bibliographic services could provide leadership in resolving.

For example, many electronic resources raise multiple-version issues, which are outside the scope of the

ISBDs. There are at least two aspects to consider. Many Internet resources are exact or related digital expressions of works in other formats. And, different digital manifestations of the same work are not uncommon. As a practical matter, how shall these electronic versions be represented in our catalogues? This phenomenon, though not unique to electronic material, is pervasive in the world of Web resources. Should each version be given its own separate bibliographic description – at the cost of convenience to the user, who normally would prefer to find all representations of a given work within a single display? Or, should the bibliographic data for all versions be combined in a single display – usually at the cost of diminished identification of bibliographic features of the separate versions that in turn decreases potential for re-use of cataloguing records? Since the ISBDs mostly provide rules for describing publications and objects at the manifestation level, national bibliographic agencies need to develop policies advising when to create separate bibliographic records and when to create a single record for resources available in two or more versions.

As another example of where national libraries and bibliographic agencies for taking a lead role in promoting standards comes about as a result of the proliferation of structures for metadata has emerged in recent years, perhaps the best known being the Dublin Core. What is characteristic to the various schemes is that they provide a structure for housing information about resources but, they offer little direction for recording data. "No bibliographic database of any significant size," as Michael Gorman has pointed out, "could possibly work if filled with Dublin Core records containing random data without vocabulary control and standard presentation."[7] The developers of metadata formats have focused on creators and distributors of electronic documents as their users, believing quite rightly that authors and publishers are not likely to be much interested in the complexities of bibliographic description, standardized access, and subject analysis. It is, therefore, opportune for the national bibliographic agency to take the lead in promoting the values of standards, especially in relation to discovery, retrieval and display of information. At the Library of Congress, we believe a first step in this direction consists of developing and disseminating a statement of basic principles to explain clearly and convincingly why there is cost-benefit from the work we do and to develop this statement with input from the metadata community. We expect to have the document ready to present publicly, perhaps at a conference, by summer 2003. Admittedly, this initiative is most likely of an educational thrust, but hopefully it will produce some tangible benefits longer term. Meanwhile, national bibliographic agencies well might want to register the emerging schema and, where possible, provide mappings from their metadata elements to fields in our established standards and formats.

Meanwhile, as Gorman goes on to note, metadata records like those called for by the Dublin Core would obviously gain value to the extent that at least some of the data elements are provided in accordance with bibliographic standards and include normalized controlled information and therefore with the benefit of professional expertise. The most obvious candidates are the fields basic to identification (such as title) and retrieval (for example, authors and subjects), data requiring librarian skills. Although such enriched records would cost more than what might have been anticipated by the creators of the Dublin Core, they would be less expensive than fully standardized records. Gorman then offers a bibliographic model for bibliographic control of remote resource that I would commend to national bibliographic agencies. Visually, the model appears as a pyramid, with a relatively small number of fully described records at the top level, with a larger number of enriched Dublin Core records in the middle followed by a layer containing Dublin Core-like records without controlled data elements, and finally a bottom layer comprising electronic resources that were not considered worthy of bibliographic control and therefore retrievable only by means of search engines.[8] The role of the national library would be to establish which electronic resources are appropriate for each of the bibliographic levels and to develop cost-effective work-flows for processing them.

The virtues of this model reflect a realistic responsiveness to the challenges of Web resources, both as to

BEST COPY AVAILABLE

their quantity and their informational value. But, other possibilities also exist to help national bibliographic services cope with the vast numbers of online resources. For example, they could develop or encourage others to develop metadata authoring tools to encourage creators of Web material to incorporate usable metadata in their products. Here the target industries include those who produce word processors, HTML editing tools, image creation and manipulation tools, and multimedia production tools. National libraries can also promote development of software to assist with creating and maintaining bibliographic records. For example, OCLC's Cooperative Online Resource Catalog (CORC) incorporates such functionality. Indeed, CORC exemplifies a bibliographic system that provides user-friendly interfaces in which persons creating catalogue records (in either MARC or Dublin Core format) are assisted by a program that proposes data for bibliographic fields using the content of the resources being catalogued. This kind of front-end to the Web could save cataloguer time and enable an increase in the number of resources listed in the national bibliography.

National libraries are also well situated to foster development of software that can detect significant changes in content of those resources catalogued and alert the agency to the need for record updates. In the U. S., as part of its Action Plan on Bibliographic Control of Web Resources, the Library of Congress has initiated a work item that is expected to result in development of this functionality.[9] The Library has also convened a working group to develop specifications for software that would facilitate maintenance of records for titles contained in aggregator databases enabling capture of accurate information about the volumes and dates of coverage of individual titles. If built into our cataloguing workflows, these tools would enable bibliographic services to improve the quality of their records for remote electronic resources.

In another strategic move, national libraries could seek opportunities to re-purpose the information generated by creators and marketers of online resources, just as they often do now in relation to the book trade. Regina Reynolds in her thought-provoking paper "Partnerships to Mine Unexploited Sources of Metadata" pointed out that in this area there is no single blueprint for establishing these partnerships. As she points out:

> To realize fully the potential of...metadata-based catalogue records, new partnerships and new sources of cataloguing data have to be explored and exploited. Metadata created in association with existing identifiers such as the ISBN and ISSN, and metadata planned to support emerging identifiers such as the Digital Object Identifier (DOI) and the developing identifier, the ISTC (International Standard Text Code), are potential sources of bibliographic data which libraries can convert, or convert and enhance to produce MARC records. Nonidentifier-based publisher registration procedures such as CIP, Copyright, and others might also yield useful data. As all of these registration procedures are increasingly completed electronically, they yield data which are highly manipulable, enhanceable, and convertible.[10]

Yet another opportunity for national bibliographic services to meet the challenges posed by online resources resides in development and utilization of staff. In her paper on "Redesign of Library Workflows: Experimental Model for Electronic Resource Description,"[11] Karen Calhoun argues that the highly centralized model for cataloguing library materials so characteristic of most libraries needs to give way to an "iterative, collaborative, and broadly distributed model". This concept values a team-based work organization, bringing together selectors, public services librarians, and cataloguers into the record creation process. Whereas in the traditional division of labor, the selector chooses resources, the cataloguer describes them, and the reference staff service them to the public, bibliographic control of remote electronic resources well may benefit from cross-functional collaboration among the selectors, cataloguers, and reference staff. Web resources lend themselves to this distributed model much more effectively than do physical objects. And, at the same time, national bibliographic agencies have the

further opportunity of promoting the use and understanding of metadata standards for describing and managing digital resources not only by their own staff but elsewhere among their constituents, particularly as concerns partnerships and cooperative arrangements whereby others contribute to the development of the national bibliographic database. Referring once again to the recommendations of the International Conference on National Bibliographic Services, we are reminded that: "National bibliographic agencies should be proactive in promoting new bibliographic standards...including holding seminars and training courses in order to ensure that both professionals and end-users are familiar with the new practices."[12]

Beyond the bibliographic problems per se are a host of other issues that national libraries are finding it necessary to address. These include copyright and legal deposit concerns. But, of great significance, is archiving and preservation of the electronic resources included in the national bibliography as well as persistent access to them. Toward the end of his presentation on "Bibliographic Control or Chaos," which has provided the basis for this workshop, Michael Gorman asked: "Supposing we solve all the problems of bibliographic standardization...what is the point if the resources identified and catalogued are not preserved?"[13] Fortunately several of today's speakers will address this question in the course of their remarks.

I would like to conclude this brief survey of the challenges confronting national bibliographic services in achieving bibliographic control of electronic resources by sharing the observations of Nancy Cline, Harvard College Librarian:

> [A]s we look to the new century, we must shape an information environment that has sustainable systems of access to enduring information resources so that users, now and in the future, can rely on them with confidence. Defining this future calls for new combinations of talent and expertise, for short- and long-term collaborations, and for experimentation and risk taking in order to develop the best strategies for managing the rapidly expanding amounts of digital information. [14]

---

1. *ISBD (NBM): International Standard Bibliographic Description for Non-Book Materials.* Recommended by the Working Group on the International Standard BibliographicDescription for Non-Book Materials set up by the IFLA Committee on Cataloguing. 1977.

2.*ISBD (CF): International Standard Bibliographic Description for Computer Files.* Recommended by the Working Group on the International Standard Bibliographic Description for Computer Files set up by the IFLA Committee on Cataloguing. 1990.

3.*ISBD (ER): International Standard Bibliographic Description for Electronic Resources.* Revised from the ISBD(CF): International Standard Bibliographic Description for Computer Files. Recommended by the ISBD(CF) Review Group. 1997.

4.Available at: <http://www.ifla.org/VI/3/icnbs/fina.htm>.

5.For further information, see the OCLC Web Characterization Project's home-page at:<http://wcp.oclc.org/ >.

6.Available at: <http://www.ifla.org/VI/3/icnbs/fina.htm>. Recommendation no. 14.

7.Gorman, Michael, "Bibliographic Control or Chaos: An Agenda for National Bibliographic Services in the 21st Century." IFLA Journal 27 5/6: pp. 312.

8. *Ibid.*

9. The LC Action Plan is available at: <http://lcweb.loc.gov/catdir/bibcontrol/actionplan.html >. The items described here are located at 4.1 and 4.2.

10. Reynolds, Regina Romano, "Partnerships to Mine Unexploited Sources of Metadata," in: *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000.* p. 439.

11.Calhoun, Karen, "Redesign of Library Workflows: Experimental Model for Electronic Resource Description,"in: *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000.* pp. 357-376.

12.Available at: <http://www.ifla.org/VI/3/icnbs/fina.htm>. Recommendation no. 20.

13. Gorman, *op. cit.*, p. 313.

14. Cline, Nancy M. "Virtual Continuity: The Challenge for Research Libraries Today." *Educause Review.* May/June 2000, p. 28.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").